2012 10th IAPR International Workshop on Document Analysis Systems

Koios++: A Query-Answering System for Handwritten Input

Marcus Liwicki*, Björn Forcher*†, Philipp Jaeger*†, Andreas Dengel*†
* *German Research Center for AI (DFKI)*
Knowledge Management Department,
Kaiserslautern, Germany
{firstname.lastname}@dfki.de
† *Knowledge-Based Systems Group,*
Department of Computer Science, University of Kaiserslautern,
P.O. Box 3049, 67653 Kaiserslautern,Germany

***Abstract*—In this paper we propose *KOIOS*++, which automatically processes natural language queries provided by handwritten input. The system integrates several recent achievements in the area of handwriting recognition, natural language processing, information retrieval, and human computer interaction. It uses a knowledge base described by the resource description framework (RDF). Our generic approach first generates a lexicon as background information for the handwritten text recognition. After recognizing a handwritten query, several output hypotheses are sent to a natural language processing system in order to generate a structured query (SPARQL query). Subsequently, the query is applied to the given knowledge base and a result graph visualizes the retrieved information. At all stages, the user can easily adjust the intermediate results if there is any undesired outcome. The system is implemented as a web-service and therefore works for handwritten input on digital paper as well as on input on Pen-enabled interactive surfaces. Furthermore, we build on the generic RDF-representation of semantic knowledge which is also used by the linked open data (LOD) initiative. As such, our system works well in various scenarios. We have implemented prototypes for querying company knowledge bases, the DBPedia[1], the DBLP computer science bibliography[2], and a knowledge base of the DAS 2012.**

Keywords*-Handwriting recognition; Information extraction; Natural language processing; HCI; semantic search; *Touch & Write

I. Introduction

The traditional aim of handwriting recognition (HWR) is to transform handwritten text into a machine readable format, i.e., the ASCII transcription has to be generated [1], [2], [3]. Research in this area continued already for several decades and there are still many open issues. Only in specific domains commercial systems were successfully applied. For online recognition, i.e., if the time-ordered point sequence of the strokes is available, recent systems perform quite well (e.g., recognizers from Microsoft© [4] and Vision Objects©). The main reasons for improved recognition performance are (1) training on large data sets of handwritten text; (2) incorporating multi-hypothesis approaches at all recognition steps; and (3) including additional information from dictionaries or language models.

Instead of just generating the ASCII transcription of the text, we address the task of understanding the meaning of the handwritten texts. Recent advantages in knowledge management and information extraction allow to make unstructured handwritten information available as structured information in a digital representation in the computer. An approach for understanding handwritten text has been proposed in [5].

In this paper we introduce a generic system which understands a natural handwritten query and retrieves the requested information from a given knowledge base. The basis of our system, i.e., the Koios system presented in [6] applies natural language processing and generates a structured query. The structured query is then used to search in the respective knowledge base. Finally, the result is presented in form of the query graph and a list of example instances in the knowledge base.

For an intuitive interaction with the application, we implemented a user interface for the *Touch & Write* table. This table allows to switch between touching and writing on the surface just by using either the fingers (touch) or an electronic pen (write). In this user interface for *KOIOS*++ the query can be written with natural handwriting on the surface and the navigation in the system as well as the selection of the desired results can be done with intuitive touch gestures.

The main contribution of the *KOIOS*++ system is that it integrates state-of-the-art approaches from handwriting recognition and human computer interaction together with natural language processing and semantic search. This novel system provides for a natural way of formulating queries for a knowledge base.

The remainder of this paper is organized as follows. First, Section II presents the background in the area of knowledge management and knowledge representation. Second, Section III gives an overview of the whole system. Next, Section IV provides more details about the semantic search engine. Subsequently, Section V describes the user interface and the behavior of the proposed system. The

[1]http://http://de.dbpedia.org/
[2]http://www.informatik.uni-trier.de/~ley/db/

978-0-7695-4661-2/12 $26.00 © 2012 IEEE
DOI 10.1109/DAS.2012.48

evaluation is reported in Section VI. Finally, Section VII draws conclusions and gives an outlook for future work.

II. Representing Knowledge

Knowledge can be represented by frames consisting of attribute-value pairs. They are often very complex and contain several kind of information, e.g., spatiotemporal schemas for contexts of events, tokens for entities like persons, objects, or ideas. Within the spatiotemporal frames there are structural relations which define a task. Those relations might be spatial, e.g., an i-dot is always placed above the i; social, e.g., I like the handwriting style of writer "A"; temporal, e.g., the signature is written at last; etc. Several contexts can also be brought into relation with one-another by linking them with causal or temporal information.

In the Semantic Web community standards, such as the Resource Description Framework[3] (RDF) and RDF Schema[4] (RDFS) have been introduced to represent the attribute-value pairs. In RDF, concepts are interlinked with one another via binary relations. This is a way of formalizing the above mentioned ideas about schemas.

Another important issue in representing knowledge is to identify digital resources, i.e., text documents, web sites, or multimedia files, by unique URIs. A very huge movement of using URIs and formal standards is the Linked Open Data (LOD) Community Project [7]. This community tries to make the web resources human- and machine understandable by describing HTTP-URIs with RDF and interlinking data from different sources using existing description standards.

Our proposed *KOIOS++* system is designed to work with RDF and with LOD. Thereby it is lifted to a generic system which works on thousands of knowledge bases available world-wide.

III. System Overview

The *KOIOS++*system retrieves information for a natural handwritten query. In contrast to typical question answering systems [8], it does not produce a natural language answer for the result, but a structured (graph-shaped) query and a list of values which fill the questioned concepts. This is motivated by the following scenario as depicted in Fig. 1.

The user has some desired information in his mind, like the papers the author "Andreas Dengel" has written. In order to get this information, he or she formulates a natural language query. This is put into the *KOIOS++* user interface (see Section V) using handwritten text. The handwritten text is recognized and the recognition alternates are sent to the *KOIOS++* semantic search engine (see Section IV).

As reported in [9], the handwriting recognition performance can be improved by integrating domain-specific knowledge from RDF-data. Since the *KOIOS++* search is performed on a given knowledge base (presented in RDF) it is wise to use this knowledge base for improving the HWR results. Therefore we adopted the methods of [9] in our system.

By matching the recognized query against the RDF-representation of the knowledge base a structured query is finally generated. Note that this structured query visualized in the bottom right of Fig. 1 is similar to the mental model of the user (top left of Fig. 1).

The *KOIOS++* system works with any kind of digital pen device. For real-time interaction, we propose the use of a pen-enabled multi-touch device. We have implemented such a user interface using the *Touch & Write* SDK. By using intuitive touch interactions (see Section V) the user can adjust the query to his or her own preference. A list of values for the questioned concepts finally contains the desired information. A more detailed description of the used methodology appears in the next two sections.

IV. The Koios Semantic Search Engine

KOIOS++ is based on a semantic search engine that enables keyword-based search on graph-shaped RDF data. *KOIOS++* first computes a set of relevant SPARQL queries allowing users to select the most appropriate ones. Consequently, a selected query is used to search in a respective triple store (knowledge base). The main advantage is that users do not need explicit knowledge about the query syntax or the underlying ontologies. The presented approach is based on the work of [10] and [11] (an illustration of the applied processes is given on the right side in Fig. 1).

To achieve a scalable search, the input data is preprocessed obtaining two data indices. The *keyword index* is used to realize a mapping from terms to resources of the RDF data, for instance, the label *Andreas Dengel* can be mapped to the instance *dblp:AndreasDengel*. Note that the index does not only contain labels of resources, it is enriched with synonyms for various resources, so that the class *dblp:Article* can be found by using terms like *paper* or *contribution*. In addition, interrogatives can be used to find upper classes, for example, the word *who* is mapped to the class *dblp:Person* and to the object property *dblp:isAuthorOf*.

For an optimized graph search, a *graph index* is derived which represents a kind of summary of the input data containing structural patterns, only. This is called *summary graph*. The summary graph contains only classes and object properties as illustrated in the bottom right corner of Fig 1. As it does not contain instances (of concepts) or literal values the summary graph is very small enabling a performant bidirectional search.

For computing SPARQL queries, the natural language query is first disassembled into its constituent parts. It is some kind of syntactic analysis in which keywords were identified that belong together. For instance, the natural language query *Andreas Dengel is author of which articles?* is separated into the parts *Andreas Dengel*, *is author of*,

[3] http://www.w3.org/RDF/
[4] http://www.w3.org/TR/rdf-schema/

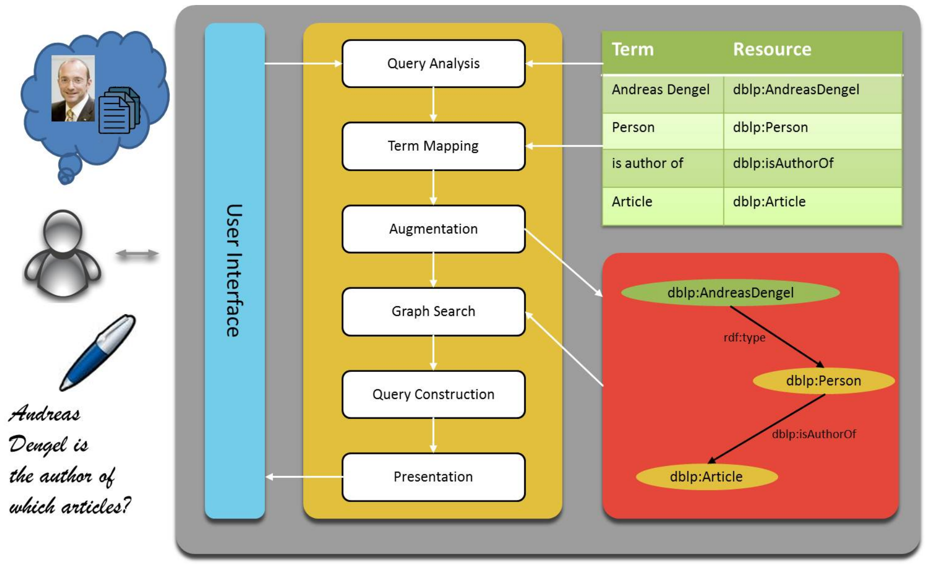

Figure 1: Overview of the *KOIOS*++ System

which and *articles*. Fur this purpose, the *Query Analysis* computes a set of useful word combinations of the natural language query, rates the combinations by means of the term index and selects the best combination for further processing. The result is a set of k terms $T = \{t_1, ..., t_k\}$, whereas $t_i \in T$ is either a single word or a multiword expression.

In the second step, the word combinations are mapped to elements of the input RDF data including the mentioned extensions. As illustrated in Fig. 1 the words *Andreas Dengel* are mapped to the RDF resource *dblp:Andreas Dengel*, whereas the interrogative *which* is mapped to the resource *dblp:Publication* that is an upper concept for the classes *dblp:Article* and *dblp:Inproceedings* in the DBLP ontology. In general, a term $t_i \in T$ is mapped to a set of n resources $R_i = \{r_{i_1}, ..., r_{i_n}\}$. As follows, there are k sets of resources $R_1, ..., R_k$ that are used for further processing.

In the next step, the summary graph is augmented by all resources of the previous step that are not contained therein. As mentioned above, this concerns literal values and instances of the base RDF data, such as the resource *dblp:AndreasDengel*. The *extension point*, i.e., the instanciated concept, is included in the term index, and thus, the resource *dblp:AndreasDengel* is attached to the summary graph node *dblp:Person*.

After the augmentation step, a graph search is performed on the augmented summary graph. The goal of this process step is to find a connection between exactly one resource of each set of resources $R_1, ..., R_k$. The outcome is a set of weighted subgraphs $S_1, ..., S_m$, whose size can be restricted by an upper weight limit and a general time limit. It would be much too far at this point to describe the weighting in detail. The interested reader may find further information in the work of Tran et al. as presented in [11]

The last two steps are straightforward. For each subgraph $S_1, ..., S_m$, a conjunctive query is constructed by mapping the graph elements to SPARQL query elements (*Query construction* step). Finally, the constructed queries are presented to the user. Note that the presentation of the search results is an important issue. Search engines should always include some hint why a particular result is relevant with respect to the keywords. Google[5], for instance, shows the title, some text snippets, and the URI of retrieved documents and highlights keywords contained therein. This information represents some kind of explanation or justification which describes a connection between search and result. In the end, it can help the user to open only relevant documents.

In case of *KOIOS*++, explaining search results for end-users is a bit more complicated. The justification has to contain the information how keywords are mapped to RDF resources and how the resources are connected. For this purpose, the UI presents the keyword mapping as well as the SPARQL query in a graph based form with standard icons as presented in Fig. 4 or Fig. 5.

V. The *KOIOS*++User Interface

Usability is an important aspect for almost all applications. Much effort have been made in Human-Computer Interaction (HCI) research.

In order to make *KOIOS*++ accessible in a natural way, we allow for handwritten input to formulate the queries. This allows users to use their own language when posting a query. The handwritten input is directly visualized in the *KOIOS*++user interface (UI) as depicted on top of Figure 2.

[5]http://www.google.de/

Figure 2: Handwritten input in the *KOIOS++* interface

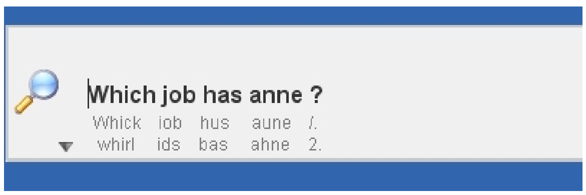

Figure 3: Visualization of the text recognition alternates

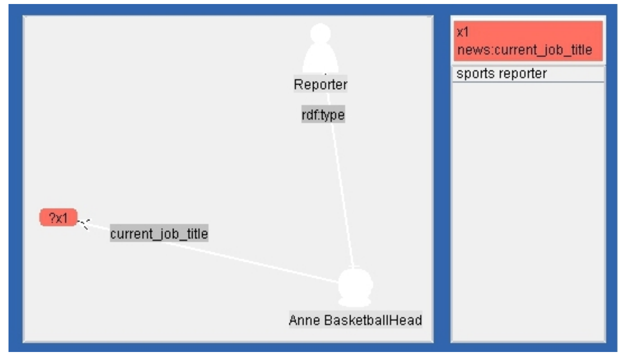

Figure 4: Retrieval result for the natural language query "Which jobs has Anne?"

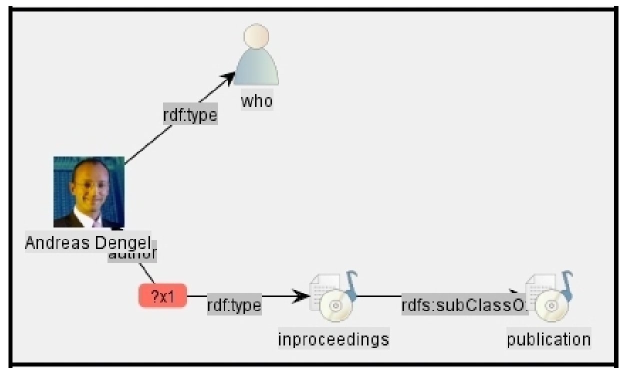

Figure 5: Retrieval result for the query "Andreas Dengel Publications"

The handwritten input is processed by the Vision Objects© handwriting recognition engine in order to generate several recognition alternates. Based on the methods of [9] and [5] the semantic search engine processes the recognition directly. In the case of recognition errors the user can simply correct the mistakes by clicking on the correct recognition alternate (see Fig. 3). This makes the interaction with our system very fast and intuitive.

As mentioned in the introduction, a further contribution is an implementation of the *KOIOS++* system on the *Touch & Write* table. The *Touch & Write* table [12] is a multi-touch tabletop solution which also integrates pen input. In particular, integrates the Anoto technology[6] on top of the usual frustrated total internal reflection (FTIR) proposed by Han [13]. Note that the integration of pen-input in multi-touch environments is quite recent achievement, since most environments lack in a way of intuitively switching between drawing and writing (Examples are the Microsoft Surface[7] or DiamondTouch [14]).

In the *Touch & Write* UI for *KOIOS++* the user can simply use the pen to write the query into the corresponding field. All further steps can be intuitively performed by touching the corresponding controls and moving the visualized concepts.

VI. Application Evaluation

Generally it is not easy to assess the performance of query answering systems because different users might expect different results when performing the same query. The main reason for this issue is the semantic gap. Despite these problems, we have investigated the behavior of the system when asking several queries. The results for two queries are depicted in Figs. 4 and 5, respectively. While the results for the first query (formulated as a complete question) is a quite useful visualization, the second query produces some output which might have been expected differently. Therefore, as stated above, the user might chose between several query representations (by browsing through several possibilities) in order to select the desired one.

More experiments have already been reported in [11]. The authors of [11] asked test persons (colleagues) to provide keyword queries along with a detailed specification in natural language of the underlying information need. 12 test persons participated in the experiment, resulting in 30 different queries for DBLP. For assessing the effectiveness of the semantic search and weighted queries, respectively, a standard IR metric called Reciprocal Rank (RR) was used in. Finally, they retrieved very satisfiable results [11].

Regarding the handwriting recognition, our previous research revealed that the combination of handwriting recognition with an information extraction engine improves the

[6]Anoto pen http://www.anoto.com/: Last accessed 04/02/2010
[7] http://www.microsoft.com/surface: Last accessed 05/22/2010

recognition results. In [5] a performance similarly to the performance on ASCII-text has been achieved.

VII. Conclusions

In this paper we presented *KOIOS++* an application which successfully integrates recent approaches in handwriting recognition, human computer interaction, natural language processing, and information extraction. In our scenario handwritten queries can be used to retrieve information from a structured knowledge base.

To the best of the authors' knowledge a query answering system for handwritten text is presented for the first time in the literature. The abstract nature of our application allows for querying any resources available in RDF, such as any LOD-resource.

We have performed an analysis of several queries in different domains. While the results of several queries seem to be convincing, we plan to perform a complete usability study of the overall *KOIOS++* system in the future. However, generally an evaluation of such complex systems is not an easy task. In order to show the effectiveness of our system, we plan to present it as a demonstration during the DAS 2012.

Acknowledgment

This work was funded by the German Federal Ministry of Education and Research (BMBF) in the EMERGENT project under grant nr. 01IC10S01. We would like to thank Markus Weber and Daniel Gröger for their technical support during development of the application.

References

[1] H. Bunke, "Recognition of cursive Roman handwriting – past present and future," in *Proc. 7th Int. Conf. on Document Analysis and Recognition*, vol. 1, 2003, pp. 448–459.

[2] R. Plamondon and S. N. Srihari, "On-line and off-line handwriting recognition: a comprehensive survey," *IEEE Trans. Pattern Analysis and Machine Intelligence*, vol. 22, no. 1, pp. 63–84, 2000.

[3] A. Vinciarelli, "A survey on off-line cursive script recognition," *Pattern Recognition*, vol. 35, no. 7, pp. 1433–1446, 2002.

[4] J. A. Pittman, "Handwriting recognition: Tablet pc text input," *Computer*, vol. 40, no. 9, pp. 49–54, 2007.

[5] S. Ebert, M. Liwicki, and A. Dengel, "Ontology-based information extraction from handwritten documents," in *12th International Conference on Frontiers in Handwriting Recognition*, 2010, pp. 483–488.

[6] B. Forcher, M. Sintek, T. Roth-Berghofer, and A. Dengel, "Explanation-aware system design of the semantic search engine koios," in *Proceedings of the ECAI-10 workshop on Explanation-aware Computing (ExACt 2010)*, 2010.

[7] T. Heath and C. Bizer, *Linked Data: Evolving the Web into a Global Data Space*, ser. Synthesis Lectures on the Semantic Web: Theory and Technology. Morgan & Claypool, 2011.

[8] H. Helbig, *Knowledge Representation and the Semantics of Natural Language (Cognitive Technologies)*. Secaucus, NJ, USA: Springer-Verlag New York, Inc., 2005.

[9] M. Liwicki, H. M. A. Eisha, and A. Dengel, "Improving handwriting recognition by the use of semantic information," in *Int. Workshop on Document Analysis Systems*, 2010, to appear.

[10] H. Wang, K. Zhang, Q. Liu, T. Tran, and Y. Yu, "Q2semantic: A lightweight keyword interface to semantic search," in *In Proceedings of the 5th International Semantic Web Conference (ESWC'08*, 2008.

[11] D. T. Tran, H. Wang, S. Rudolph, and P. Cimiano, "Top-k exploration of query candidates for efficient keyword search on graph-shaped (rdf) data," in *Proc. of the 25th Intern. Conference on Data Engineering (ICDE'09)*, Shanghai, China, March 2009.

[12] M. Liwicki, O. Rostanin, S. M. El-Neklawy, and A. Dengel, "Touch & write: a multi-touch table with pen-input," in *9th Int. Workshop on Document Analysis Systems*, 2010, pp. 479–484.

[13] J. Y. Han, "Low-cost multi-touch sensing through frustrated total internal reflection," in *UIST '05: Proceedings of the 18th annual ACM symposium on User interface software and technology*. New York, NY, USA: ACM, 2005, pp. 115–118.

[14] P. Dietz and D. Leigh, "Diamondtouch: a multi-user touch technology," in *UIST '01: Proceedings of the 14th annual ACM symposium on User interface software and technology*. New York, NY, USA: ACM, 2001, pp. 219–226.